## Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application.

## Listing of Claims:

1-18. **(Cancelled)**

19. **(Currently Amended)** A method of processing a continuous audio stream containing human speech from a plurality of speakers related to at least one particular transaction, comprising the steps of:

~~identifying a known speaker from among the plurality of speakers;~~

digitizing the continuous audio stream;

detecting a speaker change in the digitized audio stream;

performing a speaker recognition if a speaker change is detected; ~~and~~

determining whether a recognized speaker is a predetermined speaker; and

transcribing at least part of the continuous audio stream only if the ~~known~~ recognized speaker is ~~recognized~~ the predetermined speaker;

wherein said transcribing ~~each speaker~~ is processed using a ~~different~~ dictionary of ~~different~~ speaker-trained data trained by the speaker being transcribed.

20. **(Previously Presented)** A method according to claim 19, comprising a further step of protocolling time information for detected speaker changes.

21. **(Previously Presented)** A method according to claim 19, wherein the step of detecting a speaker change and/or the step of performing a speaker recognition is/are preceded by a further step of detecting non-speech boundaries between continuous speech segments.

22. **(Previously Presented)** A method according to claim 19, wherein the step of detecting a speaker change is accomplished by use of at least one characteristic audio feature, in particular features derived from the spectrum of the audio signal.

23. **(Previously Presented)** A method according to claim 19, wherein the step of performing a speaker recognition involves the particular steps of calculating a speaker signature from the audio stream and comparing the calculated speaker signature with at least one known speaker signature.

24. **(Previously Presented)** A method according to claim 19 for use in a speech recognition or voice control system comprising at least two speaker-specific speaker models and/or dictionaries, wherein interchanging between the at least two speaker-specific dictionaries is dependent on the detected speaker change and the corresponding recognized speaker.

25. **(Currently Amended)** A method of processing a continuous audio stream

containing human speech of a plurality of speakers related to at least one particular transaction,

comprising the steps of:

~~identifying a known speaker from among the plurality of speakers;~~

digitizing the continuous audio stream;

detecting a speaker change in the digitized audio stream;

performing a speaker recognition if a speaker change is detected; ~~and~~

determining whether a recognized speaker is a predetermined speaker;

indexing the audio stream with respect to the detected speaker change only if the ~~known~~

recognized speaker is ~~recognized~~ the predetermined speaker;

wherein ~~each speaker~~ said indexing is processed using a ~~different~~ dictionary of ~~different~~

speaker-trained data trained by the speaker being transcribed.

26. **(Previously Presented)** A method according to claim 25, comprising a further step

of protocolling time information for detected speaker changes.

27. **(Previously Presented)** A method according to claim 25, wherein the step of

detecting a speaker change and/or the step of performing a speaker recognition is/are preceded by

a further step of detecting non-speech boundaries between continuous speech segments.

28. **(Previously Presented)** A method according to claim 25, wherein the step of detecting a speaker change is accomplished by use of at least one characteristic audio feature, in particular features derived from the spectrum of the audio signal.

29. **(Previously Presented)** A method according to claim 25, wherein the step of performing a speaker recognition involves the particular steps of calculating a speaker signature from the audio stream and comparing the calculated speaker signature with at least one known speaker signature.

30. **(Previously Presented)** A method according to claim 25 for use in a speech recognition or voice control system comprising at least two speaker-specific speaker models and/or dictionaries, wherein interchanging between the at least two speaker-specific dictionaries is dependent on the detected speaker change and the corresponding recognized speaker.

31. **(Currently Amended)** An apparatus for processing a continuous audio stream containing human speech from a plurality of speakers related to at least one particular transaction, comprising:

~~a predeterminer which predetermines at least one known speaker from among the plurality of speakers;~~

a digitizer which digitizes the continuous audio stream;

a detector which detects speaker changes in the <u>digitized</u> audio stream;

a recognizer which recognizes the predetermined speaker in the audio stream; ~~and~~

a determiner which determines whether a recognized speaker is a predetermined speaker; and

an initiator which initiates transcription of at least part of the continuous audio stream only if the ~~in case of a detected~~ recognized speaker ~~change and a recognized~~ is the predetermined known speaker;

wherein said transcription ~~each speaker~~ is processed using a ~~different~~ dictionary of ~~different~~ speaker-trained data trained by the speaker being transcribed.

32. **(Previously Presented)** An apparatus according to claim 31, further comprising a detector which detects non-speech boundaries between continuous speech segments.

33. **(Previously Presented)** An apparatus according to claim 31, further comprising a scanner which automatically scans a continuous audio record, in particular a continuous audio stream recorded on a data or a signal carrier, and for detecting speaker changes in the continuous audio record.

34. **(Previously Presented)** An apparatus according to claim 31, further comprising a monitor which continuously monitors a real-time continuous audio stream and performing the steps of

digitizing the continuous audio stream;

detecting a speaker change in the digitized audio stream;

performing a speaker recognition if a speaker change is detected; and

transcribing at least part of the continuous audio stream if a predetermined speaker is recognized.

35. **(Previously Presented)** An apparatus according to claim 31, further comprising a monitor which continuously monitors a real-time continuous audio stream and performing the steps of

digitizing the continuous audio stream;

detecting a speaker change in the digitized audio stream;

performing a speaker recognition if a speaker change is detected; and

indexing the audio stream with respect to the detected speaker change if a predetermined speaker is recognized.

36. **(Previously Presented)** An apparatus according to claim 31, further comprising a logging device which protocols time information for the at least one detected speaker change.

37. **(Previously Presented)** An apparatus according to claim 31, comprising a marking device which marks at least the beginning of a detected speech segment related to a predetermined speaker.

38. **(Previously Presented)** An apparatus according to claim 31, comprising data base which stores speech signatures for at least two speakers.

39. (**Currently Amended**) An apparatus for processing a continuous audio stream containing human speech from a plurality of speakers related to at least one particular transaction, comprising:

~~a predeterminer which predetermines at least one known speaker from among the plurality of speakers;~~

a detector which detects speaker changes in the audio stream;

<u>a digitizer which digitizes the continuous audio stream;</u>

a recognizer which recognizes the predetermined speaker in the <u>digitized</u> audio stream; ~~and~~

<u>a determiner which determines whether a recognized speaker is a predetermined speaker;</u> <u>and</u>

an indexer for indexing <u>at least part of</u> the <u>continuous</u> audio stream <u>only if the recognized</u> ~~dependent on a detected~~ speaker <u>is the</u> ~~change and a recognized~~ predetermined <u>known</u> speaker;

wherein <u>said indexing</u> ~~each speaker~~ is processed using a ~~different~~ dictionary of ~~different topics~~ <u>speaker-trained data trained by the speaker being transcribed</u>.

40. (**Previously Presented**) An apparatus according to claim 39, further comprising a detector which detects non-speech boundaries between continuous speech segments.

41. **(Previously Presented)** An apparatus according to claim 39, further comprising a scanner which automatically scans a continuous audio record, in particular a continuous audio stream recorded on a data or a signal carrier, and for detecting speaker changes in the continuous audio record.

42. **(Previously Presented)** An apparatus according to claim 39, further comprising a monitor which continuously monitors a real-time continuous audio stream and performing the steps of

digitizing the continuous audio stream;

detecting a speaker change in the digitized audio stream;

performing a speaker recognition if a speaker change is detected; and

transcribing at least part of the continuous audio stream if a predetermined speaker is recognized.

43. **(Previously Presented)** An apparatus according to claim 39, further comprising a monitor which continuously monitors a real-time continuous audio stream and performing the steps of

digitizing the continuous audio stream;

detecting a speaker change in the digitized audio stream;

performing a speaker recognition if a speaker change is detected; and

indexing the audio stream with respect to the detected speaker change if a predetermined speaker is recognized.

44. **(Previously Presented)** An apparatus according to claim 39, further comprising a logging device which protocols time information for the at least one detected speaker change.

45. **(Previously Presented)** An apparatus according to claim 39, comprising a marking device which marks at least the beginning of a detected speech segment related to a predetermined speaker.

46. **(Previously Presented)** An apparatus according to claim 39, comprising data base which stores speech signatures for at least two speakers.

47. **(Cancelled)**

48. **(Currently Amended)** A program storage device readable by machine, tangibly embodying a program of instructions executable by the machine to perform method steps for processing a continuous audio stream containing human speech from a plurality of speakers related to at least one particular transaction, said method comprising the steps of:

identifying a known speaker from among the plurality of speakers;

digitizing the continuous audio stream;

detecting a speaker change in the digitized audio stream;

performing a speaker recognition if a speaker change is detected; and

determining whether a recognized speaker is a predetermined speaker; and

transcribing at least part of the continuous audio stream <u>only</u> if the ~~known~~ <u>recognized</u> speaker is ~~recognized~~ <u>the predetermined speaker</u>;

wherein <u>said transcribing</u> ~~each speaker~~ is processed using a ~~different~~ dictionary of ~~different~~ speaker-trained data <u>trained by the speaker being transcribed</u>.

49. **(Currently Amended)** A program storage device readable by machine, tangibly embodying a program of instructions executable by the machine to perform method steps for processing a continuous audio stream containing human speech from a plurality of speakers related to at least one particular transaction, said method comprising the steps of:

~~identifying a known speaker from among the plurality of speakers;~~

digitizing the continuous audio stream;

detecting a speaker change in the digitized audio stream;

performing a speaker recognition if a speaker change is detected; ~~and~~

<u>determining whether a recognized speaker is a predetermined speaker;</u>

indexing the audio stream with respect to the detected speaker change <u>only</u> if the ~~known~~ <u>recognized</u> speaker is ~~recognized~~ <u>the predetermined speaker</u>;

wherein ~~each speaker~~ <u>said indexing</u> is processed using a ~~different~~ dictionary of ~~different~~ speaker-trained data <u>trained by the speaker being transcribed</u>.